EXHIBIT 8

Fiber-Optic Communication Systems

SECOND EDITION

Govind P. Agrawal

Wiley Series in Microwave and Optical Engineering   Kai Chang, Series Editor

Fiber-Optic Communication Systems

Second Edition

GOVIND P. AGRAWAL
The Institute of Optics
University of Rochester
Rochester, NY

A WILEY-INTERSCIENCE PUBLICATION
JOHN WILEY & SONS, INC.
NEW YORK / CHICHESTER / WEINHEIM / BRISBANE / SINGAPORE / TORONTO

This text is printed on acid-free paper.

Library of Congress Cataloging in Publication Data:

Agrawal, G. P. (Govind P.), 1951-
Fiber-optic communication systems / Govind P. Agrawal. — 2nd ed.
p. cm. — (Wiley series in microwave and optical engineering)
"A Wiley-Interscience publication."
Includes index.
ISBN 0-471-17540-4 (cloth : alk. paper)
1. Optical communications. 2. Fiber optics. I. Title.
II. Series.
TK5103.59.A37 1997
621.382'75—dc21 97-4040

Printed in the United States of America

10 9 8 7 6 5 4 3

OYSTERDEFS2_002299

if the optical pulse spreads during transmission. The NRZ format is often used in practice because of a smaller signal bandwidth associated with it. The RZ format is required for soliton communication systems (discussed in Chapter 10).

An important issue is related to the choice of the physical variable that is modulated to encode the data on the optical carrier. The optical carrier wave before modulation is of the form

$$\mathbf{E}(t) = \hat{\mathbf{e}} A \cos(\omega_0 t + \phi), \tag{1.2.5}$$

where $\mathbf{E}$ is the electric field vector, $\hat{\mathbf{e}}$ is the polarization unit vector, A is the amplitude, ω_0 is the carrier frequency, and ϕ is the phase. The spatial dependence of $\mathbf{E}$ is suppressed for simplicity of notation. One may choose to modulate the amplitude A, the frequency ω_0, or the phase ϕ. In the case of analog modulation, the three modulation choices are known as *amplitude modulation* (AM), *frequency modulation* (FM), and *phase modulation* (PM). The same modulation techniques can be applied in the digital case and are called *amplitude-shift keying* (ASK), *frequency-shift keying* (FSK) and *phase-shift keying* (PSK), depending on whether the amplitude, frequency, or phase of the carrier wave is shifted between the two levels of a binary digital signal. The simplest technique consists of simply changing the signal intensity between two levels, one of which is set to zero, and is often called *on–off keying* (OOK) to reflect the on–off nature of the resulting optical signal. Most digital lightwave systems employ OOK in combination with PCM.

1.3 OPTICAL COMMUNICATION SYSTEMS

As mentioned earlier, optical communication systems differ in principle from microwave systems only in the frequency range of the carrier wave used to carry the information. The optical carrier frequency is typically ~ 100 THz, in contrast with the microwave carrier frequencies, $\sim$ 1–10 GHz. An increase in the information capacity of optical communication systems by a factor of $\sim 10{,}000$ is expected simply because of such high carrier frequencies used for lightwave systems. This increase can be understood by noting that the bandwidth of the modulated carrier can be up to a few percent of the carrier frequency. Taking, for illustration, 1% as the limiting value, optical communication systems have the potential of carrying information at bit rates ~ 1 Tb/s. It is this enormous potential bandwidth of optical communication systems that is the driving force behind the worldwide development and deployment of lightwave systems. Current state-of-the-art systems operate at bit rates ~ 10 Gb/s, indicating that there is considerable room for improvement.

Figure 1.8 shows a generic block diagram of an optical communication system. It consists of a transmitter, a communication channel, and a receiver, the three elements common to all communication systems. Optical communication systems can be classified into two broad categories: *guided* and *unguided*. As the name implies, in the case of guided lightwave systems, the optical beam emitted by the transmitter remains spatially confined. This is achieved by using optical

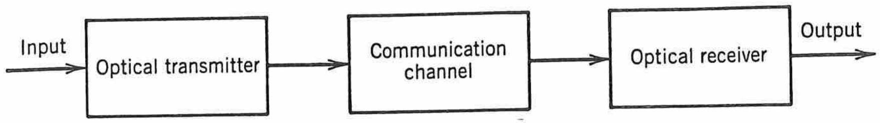

Figure 1.8 Generic optical communication system.

fibers, as discussed in Chapter 2. Since all guided optical communication systems currently use optical fibers, the commonly used term for them is fiber-optic communication systems. The term *lightwave system* is also sometimes used for fiber-optic communication systems, although it should generally include both guided and unguided systems.

In the case of unguided optical communication systems, the optical beam emitted by the transmitter spreads in space, similar to the spreading of microwaves. However, unguided optical systems are less suitable for broadcasting applications than microwave systems because optical beams spread mainly in the forward direction (as a result of their short wavelength). Their use generally requires accurate pointing between the transmitter and the receiver. In the case of terrestrial propagation, the signal in unguided systems can deteriorate considerably by scattering within the atmosphere. This problem, of course, disappears in *free-space communications* above the earth atmosphere (e.g., intersatellite communications). Although free-space optical communication systems are needed for certain applications and have been studied extensively [62], most terrestrial applications make use of *fiber-optic communication systems.* This book does not consider unguided optical communication systems.

The application of optical fiber communications is in general possible in any area that requires transfer of information from one place to another. However, fiber-optic communication systems have been developed mostly for telecommunications applications. This is understandable in view of the existing worldwide telephone networks which are used to transmit not only voice signals but also computer data and fax messages. The telecommunication applications can be broadly classified into two categories, *long-haul* and *short-haul*, depending on whether the optical signal is transmitted over relatively long or short distances compared with typical intercity distances ($\sim$ 100 km). Long-haul telecommunication systems require high-capacity trunk lines and benefit most by the use of fiber-optic lightwave systems. Indeed, the technology behind optical fiber communication is often driven by long-haul applications. Each successive generation of lightwave systems is capable of operating at higher bit rates and over longer distances. Periodic regeneration of the optical signal by using repeaters is still required for most long-haul systems. However, more than an order-of-magnitude increase in both the repeater spacing and the bit rate compared with those of coaxial systems has made the use of lightwave systems very attractive for long-haul applications. Furthermore, transmission distances of thousands of kilometers can be realized by using optical amplifiers. As shown in Fig. 1.3, a large number of transoceanic lightwave systems have already been installed to create an international fiber-optic network.

Short-haul telecommunication applications cover intracity and local-loop traffic. Such systems typically operate at low bit rates over distances of less than 10 km. The use of single-channel lightwave systems for such applications is not very cost-effective, and multichannel networks with multiple services should be considered. The concept of a *broadband integrated-services digital network* requires a high-capacity communication system capable of carrying multiple services. The ATM technology also demands high bandwidths [61]. Only fiber-optic communication systems are likely to meet such wideband distribution requirements. Multichannel lightwave systems and their applications in local-area networks are discussed in Chapter 7.

1.4 LIGHTWAVE SYSTEM COMPONENTS

The generic block diagram of Fig. 1.8 applies to a fiber-optic communication system, the only difference being that the communication channel is an optical fiber cable. The other two components, the optical transmitter and the optical receiver, are designed to meet the needs of such a specific communication channel. In this section we discuss the general issues related to the role of optical fiber as a communication channel and to the design of transmitters and receivers. The objective is to provide an introductory overview, as the three components are discussed in detail in Chapters 2–4.

1.4.1 Optical Fibers as a Communication Channel

The role of communication channel is to transport the optical signal from transmitter to receiver without distorting it. Most lightwave systems use optical fibers as the communication channel because fibers can transmit light with a relatively small amount of power loss. In Chapter 2 we discuss the properties of optical fibers in detail. Fiber loss is, of course, an important design issue, as it determines directly the repeater spacing of a long-haul lightwave system. Another important design issue is *fiber dispersion*, which leads to broadening of individual optical pulses inside the fiber. If optical pulses spread significantly outside their allocated bit slot, the transmitted signal is severely degraded. Eventually, it becomes impossible to recover the original signal with high accuracy. The problem is most severe in the case of multimode fibers, since pulses spread rapidly (typically at a rate of ~ 10 ns/km) because of different speeds associated with different fiber modes. It is for this reason that most optical communication systems use single-mode fibers. *Material dispersion* (related to the frequency dependence of the refractive index) still leads to pulse broadening (typically < 0.1 ns/km), but it is small enough to be acceptable for most applications and can be reduced further by controlling the spectral width of the optical source. Nevertheless, as discussed in Chapter 2, material dispersion sets the ultimate limit on the bit rate and the transmission distance of fiber-optic communication systems.

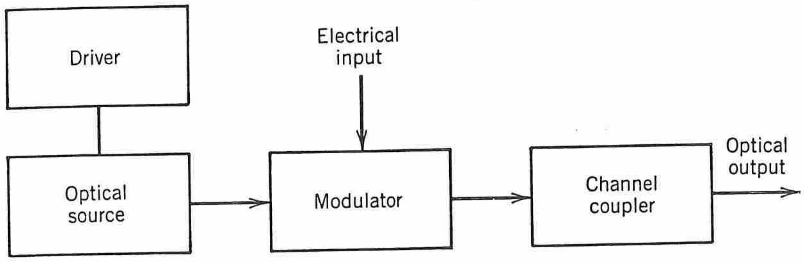

Figure 1.9 Components of an optical transmitter.

1.4.2 Optical Transmitters

The role of an *optical transmitter* is to convert the electrical signal into optical form and to launch the resulting optical signal into the optical fiber. Figure 1.9 shows the block diagram of an optical transmitter. It consists of an optical source, a modulator, and a channel coupler. Semiconductor lasers or light-emitting diodes are used as optical sources because of their compatibility with the optical-fiber communication channel; both are discussed in detail in Chapter 3. The optical signal is generated by modulating the optical carrier wave. Although an external modulator is sometimes used, it can be dispensed with in most cases, since the output of a semiconductor optical source can be modulated directly by varying the injection current. Such a scheme simplifies the transmitter design and is generally cost-effective. The coupler is typically a microlens that focuses the optical signal onto the entrance plane of an optical fiber with the maximum possible efficiency.

The *launched power* is an important design parameter, as it indicates how much fiber loss can be tolerated. It is often expressed in units of dBm with 1 mW as the reference level. The general definition is (see Appendix B)

$$\text{power (dBm)} = 10\log_{10}\left(\frac{\text{power}}{1\text{ mW}}\right). \tag{1.4.1}$$

Thus, 1 mW is 0 dBm, but 1 μW corresponds to -30 dBm. The launched power is rather low (< -10 dBm) in the case of light-emitting diodes, whereas semiconductor lasers can launch powers ~ 10 dBm. Since light-emitting diodes are also limited in their modulation capabilities, most high-performance lightwave systems use semiconductor lasers as optical sources. The bit rate of optical transmitters is often limited by electronics rather than by the semiconductor laser itself. With proper design, optical transmitters can be made to operate at a bit rate of up to 20 Gb/s.

1.4.3 Optical Receivers

An *optical receiver* converts the optical signal received at the output end of the optical fiber back into the original electrical signal. Figure 1.10 shows the block diagram of an optical receiver. It consists of a coupler, a photodetector,

For the case of homodyne detection, SNR is larger by a factor of 2 and is given by SNR $= 4\eta N_p$. Section 6.4 discusses the dependence of the BER on SNR and shows how receiver sensitivity is improved by the use of coherent detection.

6.2 MODULATION FORMATS

As discussed in Section 6.1, an important advantage of using the coherent detection techniques is that both the amplitude and the phase of the received optical signal can be detected and measured. This feature opens up the possibility of sending information by modulating either the amplitude, or the phase, or the frequency of the optical carrier. In the case of digital communication systems, the three possibilities give rise to three modulation formats: *amplitude-shift keying* (ASK), *phase-shift keying* (PSK), and *frequency-shift keying* (FSK) [1]–[6]. Figure 6.2 shows schematically the three modulation formats for a specific bit pattern. In the following subsections we consider each format separately and discuss its implementation in practical lightwave systems.

6.2.1 ASK Format

The electric field associated with the optical signal can be written as [by taking the real part of Eq. (6.1.1)]

$$E_s(t) = A_s(t)\cos[\omega_0 t + \phi_s(t)]. \tag{6.2.1}$$

In the case of ASK format, the amplitude A_s is modulated while keeping ω_0 and ϕ_s constant. For binary digital modulation, A_s takes one of the two fixed values during each bit period, depending on whether "1" or "0" bit is being transmitted. In most practical situations, A_s is set to zero during transmission of "0" bits. The ASK format is then called *on–off keying* (OOK) and is identical with the modulation scheme commonly used for noncoherent (IM/DD) digital lightwave systems.

The implementation of ASK for coherent systems differs considerably from the case of the direct-detection systems discussed in Chapter 5. Whereas the optical bit stream for direct-detection systems is often generated by modulating a LED or a semiconductor laser directly, an external modulator becomes a necessity for coherent communication systems. The reason behind this necessity is related to phase changes that invariably occur when the amplitude A_s (or the power) is changed by modulating the current applied to a semiconductor laser (see Section 3.3.7). For IM/DD systems, such unintentional phase changes are not seen by the detector (as the detector responds only to the optical power) and are not of major concern except for the chirp-induced power penalty discussed in Section 5.4.4. The situation is entirely different in the case of coherent systems, where the detector response depends on the phase of the received signal. The implementation of ASK format for coherent systems requires the phase ϕ_s to remain nearly constant. This is achieved by operating the semiconductor laser continuously at a constant current and modulating its output by using an

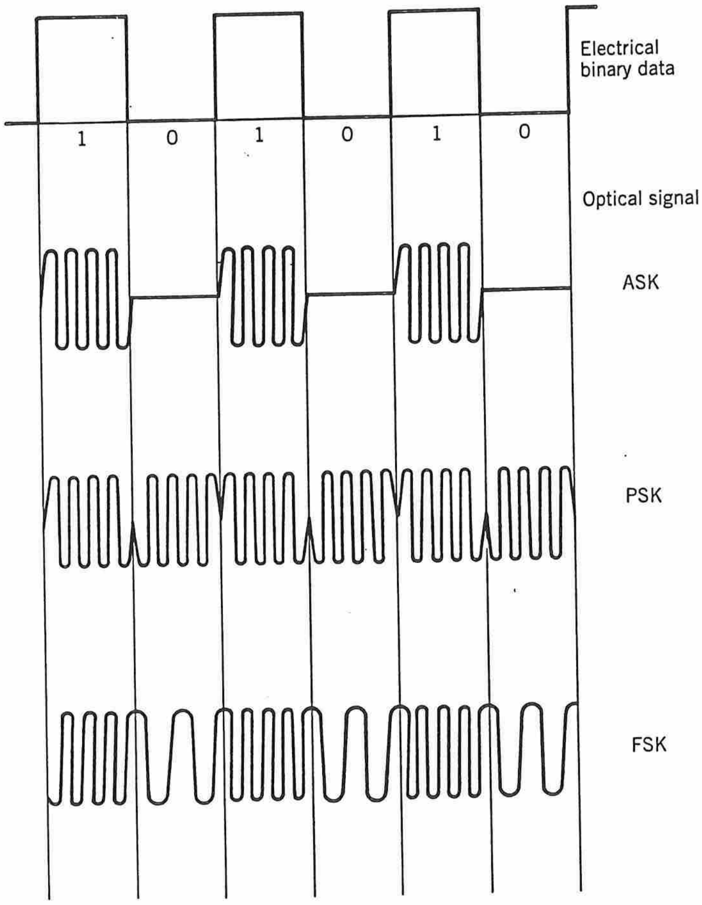

Figure 6.2 ASK, PSK, and FSK modulation formats for a specific bit pattern shown on the top. In each case rapid oscillations correspond to variations of the electromagnetic field at the optical carrier frequency.

external modulator. Since all external modulators have some insertion loss, a power penalty is generally paid whenever an external modulator is used; it can be reduced to below 1 dB for monolithically integrated modulators.

Commonly used external modulators make use of titanium-diffused $LiNbO_3$ waveguides in a Mach–Zehnder or a directional-coupler configuration [17]–[20]. The Mach–Zehnder interferometer design is shown in Fig. 6.3(a). The refractive index of electro-optic materials such as $LiNbO_3$ can be changed by applying an external voltage. In the absence of external voltage, the optical fields in the two arms of the Mach–Zehnder interferometer experience identical phase shifts and interfere constructively. The additional phase shift introduced in one of the arms through voltage-induced index changes destroys the constructive nature of the interference and reduces the transmitted intensity. In particular, no light is transmitted when the phase difference between the two arms equals

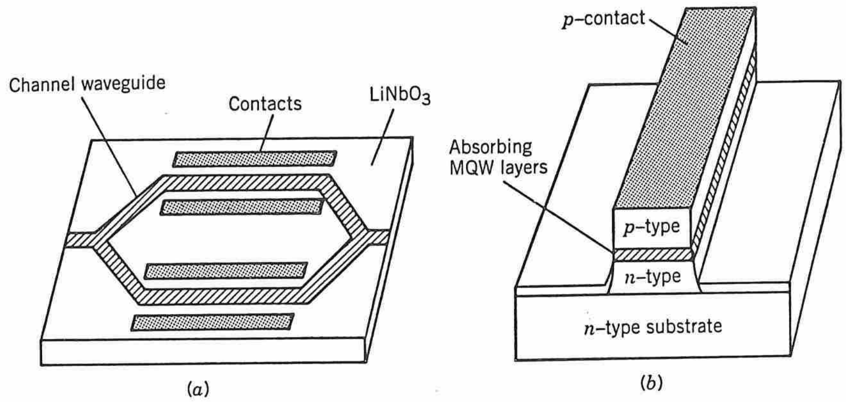

Figure 6.3 Two kinds of external modulators used for ASK: (a) $LiNbO_3$ waveguide modulator in the Mach–Zehnder configuration; (b) semiconductor waveguide modulator based on electroabsorption.

π, because of destructive interference occurring in that case. As a result, the electrical bit stream applied to the modulator produces an optical replica of the bit stream. The performance of external modulators is quantified through the on–off ratio (also called extinction ratio) and the modulation bandwidth. $LiNbO_3$ modulators provide an on–off ratio in excess of 20 and can be modulated at speeds up to 75 GHz [18], [19]. The driving voltage is typically 5 V but can be reduced to near 3 V with a suitable design [20]. Other materials can also be used to make external modulators. For example, modulators have been fabricated using electro-optic polymers, and modulation bandwidths of up to 60 GHz have been demonstrated [21]. Such modulators can be integrated monolithically with the driving circuitry [22].

Modulators can also be made using semiconductors. The basic mechanism behind the operation of semiconductor modulators is electroabsorption discussed in Section 3.4.4. All semiconductors begin to absorb incident light when photon energy exceeds the bandgap energy. Since the electroabsorption effect is stronger in multiquantum-well (MQW) structures, such a structure [see Fig. 6.3(b)] has been studied extensively [23]–[26]. Modulation bandwidths in excess of 20 GHz were demonstrated [23] as early as 1989 and later extended to 50 GHz. By 1996, polarization-insensitive 1.55-μm modulators with a large bandwidth (42 GHz) and a low operating voltage (< 2 V) have been made using a strained-MQW design [26]. An advantage of semiconductor modulators is that they can be integrated with the semiconductor laser on the same chip (see Section 3.4.4). Such monolithically integrated transmitters can operate at bit rates as high as 20 Gb/s and are available commercially.

6.2.2 PSK Format

In the case of PSK format, the optical bit stream is generated by modulating the phase ϕ_s in Eq. (6.2.1) while the amplitude A_s and the frequency ω_0 of

the optical carrier are kept constant. For binary PSK, the phase ϕ_s takes two values, commonly chosen to be 0 and π. Figure 6.2 shows the binary PSK format schematically for a specific bit pattern. An interesting aspect of the PSK format is that the optical intensity remains constant during all bits and the signal appears to have a continuous-wave (CW) form. Coherent detection is a necessity for PSK, as all information would be lost if the optical signal were detected directly without mixing it with the output of a local oscillator.

The implementation of PSK requires an external modulator capable of changing the optical phase in response to an applied voltage. The physical mechanism used by such modulators is called electrorefraction. Any electro-optic crystal with proper orientation can be used for phase modulation. A $LiNbO_3$ crystal is commonly used in practice. The design of $LiNbO_3$-based phase modulators is much simpler than that shown in Fig. 6.3(a), as a Mach–Zehnder interferometer is no longer needed; a single waveguide is used in its place. The phase shift $\delta\phi$ is related to the index change δn by the simple relation

$$\delta\phi = (2\pi/\lambda)(\delta n)l_m, \tag{6.2.2}$$

where l_m is the length over which index change is induced by the applied voltage. The index change δn is proportional to the applied voltage, which is chosen such that $\delta\phi = \pi$. A π phase shift can be imposed on the optical carrier by applying the required voltage for the duration of each "1" bit.

Semiconductors can also be used to make phase modulators, especially if a MQW structure is used, since the electrorefraction effect originating from the *quantum-confinement Stark effect* is enhanced for a quantum-well design. Such MQW phase modulators have been developed [27]–[32] and are able to operate at a bit rate of up to 20 Gb/s in the wavelength range 1.3–1.6 μm. By 1992, MQW devices had a modulation bandwidth of 20 GHz and required only 3.85 V for introducing a π phase shift when operated near 1.55 μm [27]. The operating voltage was further reduced to 2.8 V in a phase modulator based on the electro-absorption effect in a MQW waveguide [28]. However, the module exhibited an insertion loss of 8 dB because of high coupling losses. A spot-size converter is sometimes integrated with the phase modulator to reduce coupling losses [29]. The best performance is achieved when a semiconductor phase modulator is monolithically integrated within the transmitter [30]. Such transmitters are quite useful for coherent lightwave systems. Several other techniques can be used to make phase modulators [31], [32].

The use of PSK format requires that the phase of the optical carrier remain stable so that phase information can be extracted at the receiver without ambiguity. This requirement puts a stringent condition on the tolerable linewidths of the transmitter laser and the local oscillator. As discussed later in Section 6.5.1, the linewidth requirement can be somewhat relaxed by using a variant of the PSK format, known as *differential phase-shift keying* (DPSK). In the case of DPSK, information is coded by using the phase difference between two neighboring bits. For instance, if ϕ_k represents the phase of the kth bit, the phase difference $\Delta\phi = \phi_k - \phi_{k-1}$ is changed by π or 0, depending on whether kth bit

is a 1 or 0 bit. The advantage of DPSK is that the transmittal signal can be demodulated successfully as long as the carrier phase remains relatively stable over a duration of two bits.

6.2.3 FSK Format

In the case of FSK modulation, information is coded on the optical carrier by shifting the carrier frequency ω_0 itself [see Eq. (6.2.1)]. For a binary digital signal, ω_0 takes two values, $\omega_0 + \Delta\omega$ and $\omega_0 - \Delta\omega$, depending on whether a 1 or 0 bit is being transmitted. The shift $\Delta f = \Delta\omega/2\pi$ is called the *frequency deviation*. The quantity $2\Delta f$ is sometimes called *tone spacing*, as it represents the frequency spacing between 1 and 0 bits. The optical field for FSK format can be written as

$$E_s(t) = A_s \cos[(\omega_0 \pm \Delta\omega)t + \phi_s], \tag{6.2.3}$$

where + and − signs correspond to 1 and 0 bits. By noting that the argument of cosine can be written as $\omega_0 t + (\phi_s \pm \Delta\omega t)$, the FSK format can also be viewed as a kind of PSK modulation such that the carrier phase increases or decreases linearly over the bit duration.

The choice of the frequency deviation Δf depends on the available bandwidth. The total bandwidth of a FSK signal is given approximately by $2\Delta f + 2B$, where B is the bit rate [1]. When $\Delta f \gg B$, the bandwidth approaches $2\Delta f$ and is nearly independent of the bit rate. This case is often referred to as *wide-deviation* or wideband FSK. In the opposite case of $\Delta f \ll B$, called *narrow-deviation* or narrowband FSK, the bandwidth approaches $2B$. The ratio $\beta_{FM} = \Delta f/B$, called the FM index, serves to distinguish the two cases, depending on whether $\beta_{FM} \gg 1$ or $\beta_{FM} \ll 1$.

The implementation of FSK requires modulators capable of shifting the frequency of the incident optical signal. Electro-optic materials such as $LiNbO_3$ normally produce a phase shift proportional to the applied voltage. They can be used for FSK by applying a triangular voltage pulse (sawtooth-like), since a linear phase change corresponds to a frequency shift. An alternative technique makes use of Bragg scattering from acoustic waves. Such modulators are called acousto-optic modulators. Their use is somewhat cumbersome in the bulk form. However, they can be fabricated in compact form by using surface acoustic waves on a slab waveguide. A problem with the acousto-optic modulators is that a frequency shift is accompanied by a shift in the diffraction angle, making alignment difficult for large frequency shifts. Generally, the frequency shifts are limited to below 1 GHz for such techniques.

The simplest and most commonly used method for the FSK format makes use of the direct-modulation capability of semiconductor lasers. As discussed in Section 3.4, a change in the operating current of a semiconductor laser leads to changes in both the amplitude and frequency of emitted light. In the case of ASK or OOK, the frequency shift chirps the emitted optical pulse and is undesirable. But the same frequency shift can be used to advantage for the purpose of FSK. Typical values of frequency shifts are 0.1–1 GHz/mA. Therefore, only a small

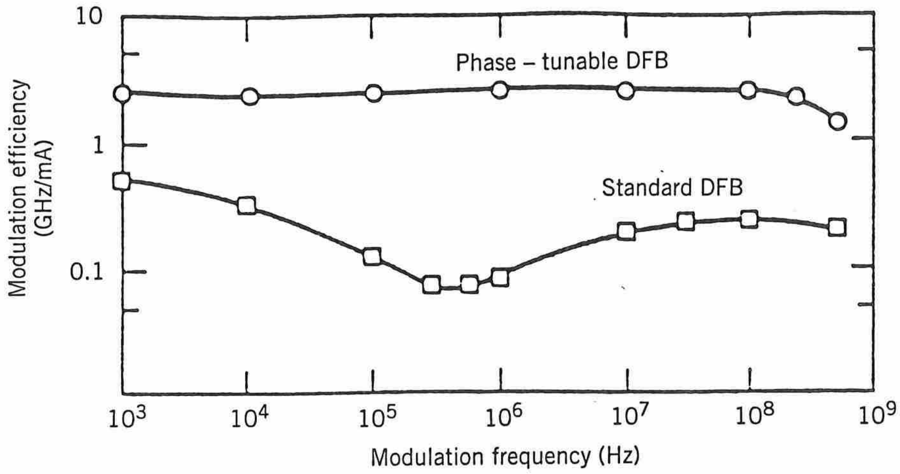

Figure 6.4 FM response of a typical DFB semiconductor laser exhibiting a dip in the frequency range 0.1–10 MHz. (After Ref. [12]. ©1988 IEEE. Reprinted with permission.)

change in the operating current ($\sim$ 1 mA) is required to produce frequency shifts of 1 GHz. Such current changes are small enough that the amplitude changes little from bit to bit.

DFB semiconductor lasers are often used for coherent transmission because of their ability to operate in a single longitudinal mode with a narrow linewidth. For the purpose of FSK, the FM response of such lasers should be flat over the entire bandwidth. Unfortunately, this is not the case in practice. Semiconductor lasers typically exhibit a dip in their FM response over the frequency range 0.1–10 MHz [33]. The reason is that two different physical phenomena contribute to the frequency shift when the device current is changed. Changes in the refractive index, responsible for the frequency shift, can occur either because of a temperature shift or because of a change in the carrier density. The thermal effects contribute only up to modulation frequencies of about 1 MHz because of a slow thermal response. The FM response decreases in the frequency range 0.1–10 MHz because the thermal contribution and the carrier-density contribution occur with opposite phases. Figure 6.4 shows the FM response of a typical DFB laser.

The nonuniformity of FM response is undesirable for FSK coherent lightwave systems, and many techniques are used to make it uniform. An equalization circuit can be used, although it generally reduces the modulation efficiency. Another technique makes use of transmission codes which reduce the low-frequency components of the data where distortion is highest. New types of multisection DFB lasers have been developed to achieve uniform FM response [34]–[40]. Figure 6.4 shows the FM response of a two-section DFB laser. It is not only uniform in the range up to about 1 GHz, but its modulation efficiency is also high. Even better performance has been realized by using three-section DFB lasers,

which are described in Section 3.3.5. Such lasers consist of an active section, a phase-shift section, and a Bragg-mirror section, each of which can be independently biased (see Fig. 3.20). Flat FM response from 100 kHz to 15 GHz was demonstrated [34] in 1990 in such a three-section DFB laser, which continued to operate in a single longitudinal mode with a narrow linewidth (< 1 MHz). By 1995, the use of gain-coupled, phase-shifted, DFB lasers has extended the range of uniform FM response from 10 kHz to 20 GHz [38]. When FSK is performed through direct modulation, the carrier phase varies continuously from bit to bit. The FSK format in that case is often referred to as continuous-phase FSK (CPFSK). When the tone spacing $2\Delta f$ is chosen to be $B/2$ ($\beta_{FM} = 1/2$), CPFSK is also called minimum-shift keying (MSK).

6.3 DEMODULATION SCHEMES

As discussed in Section 6.1, either homodyne or heterodyne detection can be used to convert the received optical signal into electrical form. In the case of homodyne detection, the optical signal is demodulated directly to the baseband. Although simple in concept, homodyne detection is difficult to implement in practice, as it requires a local oscillator whose frequency matches the carrier frequency exactly and whose phase is locked to the incoming signal. Such a demodulation scheme is called synchronous and is essential for homodyne detection. Although optical phase-locked loop have been developed for this purpose, their use is complicated in practice. Heterodyne detection simplifies the receiver design, as neither optical phase locking nor frequency matching of the local oscillator is required. However, the electrical signal is in the form of microwaves and must be demodulated from the intermediate frequency to the baseband by using techniques developed for microwave communication systems [1]–[6]. Demodulation can be carried out either synchronously or asynchronously. Asynchronous demodulation is also called incoherent in radio communication literature. In optical communication literature, the term *coherent detection* is used in a wider sense. A lightwave system is called coherent as long as it uses a local oscillator for homodyne or heterodyne mixing of the optical signal irrespective of the demodulation technique used to convert the intermediate-frequency signal to baseband frequencies. In the following subsections we describe synchronous and asynchronous demodulation schemes for the heterodyne signal.

6.3.1 Heterodyne Synchronous Demodulation

Figure 6.5 shows a block diagram of a synchronous heterodyne receiver. The current generated at the photodiode is passed through a bandpass filter (BPF) centered at the intermediate frequency ω_{IF}. The filtered current in the absence of noise can be written as [see Eq. (6.1.8)]

$$I_f(t) = I_p \cos(\omega_{IF} t - \phi), \tag{6.3.1}$$

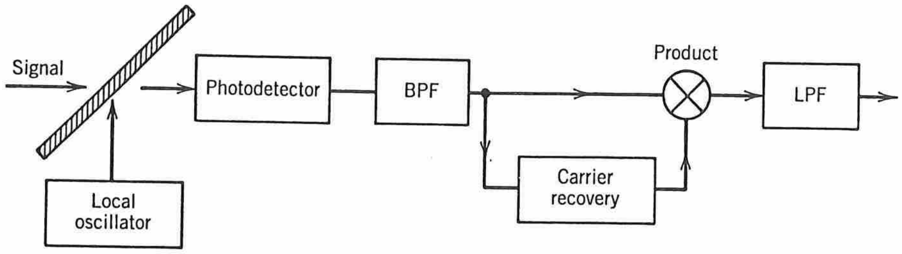

Figure 6.5 Block diagram of a synchronous heterodyne receiver.

where I_p is given by Eq. (6.1.6) and ϕ is the phase difference between the local oscillator and signal phases. The receiver noise is also filtered by the BPF. If we use the in-phase and out-of-phase quadrature components of the filtered Gaussian noise [1], the receiver noise is included through

$$I_f(t) = (I_p \cos\phi + i_c)\cos(\omega_{IF}t) + (I_p \sin\phi + i_s)\sin(\omega_{IF}t), \qquad (6.3.2)$$

where i_c and i_s are Gaussian random variables of zero mean whose variance σ^2 is given by Eq. (6.1.9). For synchronous demodulation, $I_f(t)$ is multiplied by $\cos(\omega_{IF}t)$ and filtered by a low-pass filter. The resulting baseband signal is

$$I_d = \langle I_f \cos(\omega_{IF}t)\rangle = \tfrac{1}{2}(I_p \cos\phi + i_c), \qquad (6.3.3)$$

where angle brackets denote low-pass filtering that leads to rejection of ac components oscillating at $2\omega_{IF}$. Equation (6.3.3) shows that only the in-phase noise component affects the performance of synchronous heterodyne receivers.

Synchronous demodulation requires recovery of the microwave carrier at the intermediate frequency ω_{IF}. Several electronic schemes can be used for this purpose, all requiring a kind of electrical phase-locked loop [41]. Two commonly used loops are the *squaring loop* and the *Costas loop*. A squaring loop uses a square-law device to obtain a signal of the form $\cos^2(\omega_{IF}t)$ that has a frequency component at $2\omega_{IF}$. This component can be used to generate a microwave signal at ω_{IF}.

6.3.2 Heterodyne Asynchronous Demodulation

Figure 6.6 shows a block diagram of an asynchronous heterodyne receiver. It does not require recovery of the microwave carrier at the intermediate frequency, resulting in a much simpler receiver design. The filtered signal $I_f(t)$ is converted to the baseband by using an *envelope detector*, followed by a low-pass filter. The signal received by the decision circuit is just $I_d = |I_f|$, where I_f is given by Eq. (6.3.2). It can be written as

$$I_d = |I_f| = [(I_p \cos\phi + i_c)^2 + (I_p \sin\phi + i_s)^2]^{1/2}. \qquad (6.3.4)$$

The main difference is that both the in-phase and out-of-phase quadrature components of the receiver noise affect the signal. The SNR is thus degraded compared with the case of synchronous demodulation. As discussed in Section 6.4,

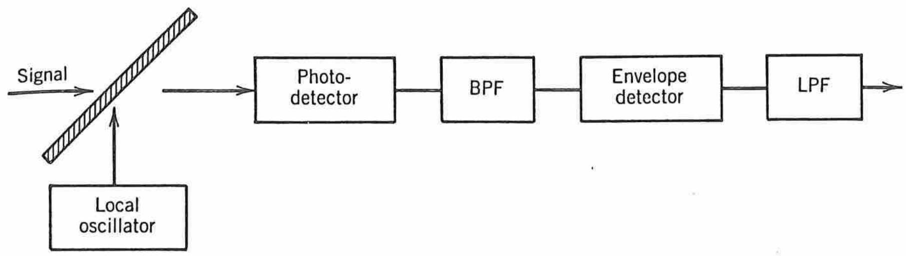

Figure 6.6 Block diagram of an asynchronous heterodyne receiver.

asynchronous receivers are less sensitive because of the SNR degradation. However, sensitivity degradation is quite small (about 0.5 dB). At the same time, the linewidth requirements for the transmitter and the local oscillator are quite modest in the case of asynchronous demodulation (see Section 6.5). For this reason, asynchronous heterodyne receivers play an important role in the design of coherent lightwave systems.

The asynchronous heterodyne receiver shown in Fig. 6.6 requires modifications when the FSK and PSK modulation formats are used. Figure 6.7 shows two demodulation schemes. The FSK *dual-filter receiver* uses two separate branches to process the 1 and 0 bits whose carrier frequencies, and hence the intermediate frequencies, are different. The scheme can be used whenever the tone spacing is much larger than the bit rates, so that the spectra of 1 and 0 bits have negligible overlap (wide-deviation FSK). The two BPFs have their center frequencies separated exactly by the tone spacing so that each BPF passes either 1 or 0 bits only. The FSK dual-filter receiver can be thought of as two ASK single-filter receivers in parallel whose outputs are combined before reaching the decision circuit. A single-filter receiver of Fig. 6.6 can be used for FSK demodulation if its bandwidth is chosen to be wide enough to pass the entire bit stream. The signal is then processed by a frequency discriminator to identify 1 and 0 bits. This scheme works well only for narrow-deviation FSK, for which tone spacing is less than or comparable to the bit rate ($\beta_{FM} \leq 1$).

Asynchronous demodulation cannot be used for the PSK format, as the phase of the transmitter laser and the local oscillator are not locked and can drift with time. However, the use of DPSK format permits asynchronous demodulation by using the delay scheme shown in Fig. 6.7(b). The idea is to multiply the received bit stream by a replica of it that has been delayed by one bit period. The resulting signal has a component of the form $\cos(\phi_k - \phi_{k-1})$, where ϕ_k is the phase of kth bit, which can be used to recover the bit pattern since information is encoded in the phase difference $\phi_k - \phi_{k-1}$. Such a scheme requires phase stability only over a few bits and can be implemented by using DFB semiconductor lasers. The delay-demodulation scheme can also be used for CPFSK. The amount of delay in that case depends on the tone spacing and is chosen such that the phase is shifted by π for the delayed signal.

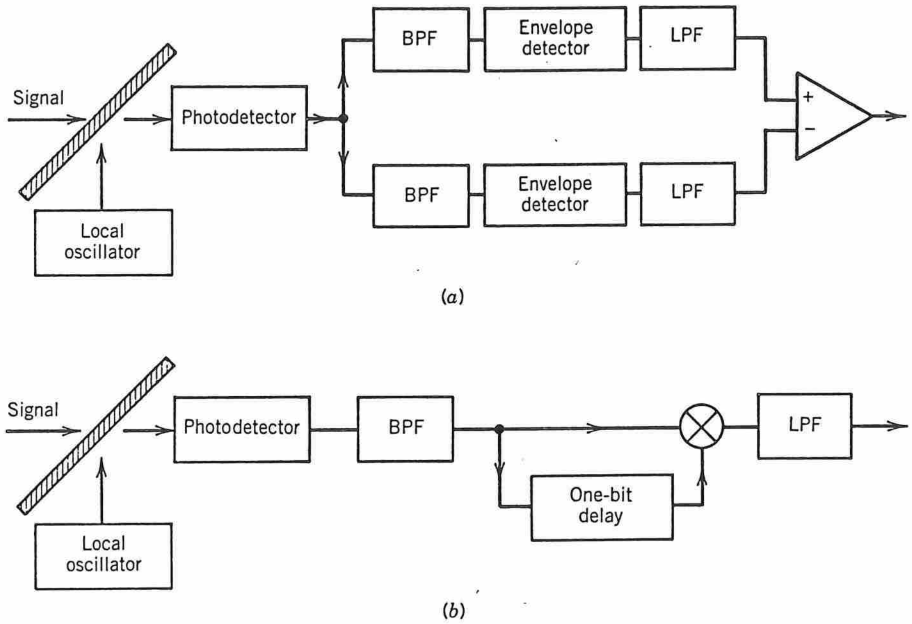

Figure 6.7 (a) Dual-filter FSK and (b) delay-demodulation DPSK asynchronous heterodyne receivers.

6.4 BIT-ERROR RATE

The preceding three sections have provided enough background material for calculating the bit-error rate (BER) of coherent lightwave systems. However, the BER, and hence the receiver sensitivity, depend on the modulation format as well as on the demodulation scheme used by the coherent receiver. The section considers each case separately.

6.4.1 Synchronous ASK Receivers

Consider first the case of heterodyne detection. The signal used by the decision circuit is given by Eq. (6.3.3). The phase ϕ generally varies randomly because of phase fluctuations associated with the transmitter laser and the local oscillator. As discussed in Section 6.5, the effect of phase fluctuations can be made negligible by using semiconductor lasers whose linewidth is a small fraction of the bit rate. Assuming this to be the case and setting $\phi = 0$ in Eq. (6.3.2), the decision signal is given by

$$I_d = \tfrac{1}{2}(I_p + i_c), \tag{6.4.1}$$

where $I_p = 2R(P_s P_{LO})^{1/2}$ from Eq. (6.1.6). I_p takes values I_1 or I_0 depending on whether 1 or 0 bit is being detected. For simplicity, consider the case $I_0 = 0$, in which no power is transmitted during the 0 bits. Except for the factor of 1/2 in Eq. (6.4.1), the situation is analogous to the case of direct detection discussed

in Section 4.5. The factor of 1/2 does not affect the BER since both the signal and the noise are reduced by the same factor, leaving the SNR unchanged. In fact, one can use the same result [Eq. (4.5.9)],

$$\mathrm{BER} = \frac{1}{2}\,\mathrm{erfc}\left(\frac{Q}{\sqrt{2}}\right), \tag{6.4.2}$$

where Q is given by Eq. (4.5.10) and can be written as

$$Q = \frac{I_1 - I_0}{\sigma_1 + \sigma_0} \approx \frac{I_1}{2\sigma_1} = \frac{1}{2}(\mathrm{SNR})^{1/2}. \tag{6.4.3}$$

In relating Q to SNR, we used $I_0 = 0$ and set $\sigma_0 \approx \sigma_1$. The latter approximation is justified for most coherent receivers whose noise is dominated by the shot noise induced by local-oscillator power and remains the same irrespective of the received signal power. Indeed, as shown in Section 6.1.4, the SNR of such receivers can be related to the number of photons received during each 1 bit by the simple relation $\mathrm{SNR} = 2\eta N_p$ [see Eq. (6.1.15)], where η is the quantum efficiency of the photodetector. Equations (6.4.2) and (6.4.3) together with the relation above provide the following expression for the BER:

$$\mathrm{BER} = \tfrac{1}{2}\,\mathrm{erfc}(\sqrt{\eta N_p/4}). \qquad \text{[ASK heterodyne]} \tag{6.4.4}$$

One can use the same method to calculate the BER for the case of ASK homodyne receivers. Equations (6.4.2) and (6.4.3) still remain applicable. However, the SNR is improved by 3 dB for the homodyne case, so that $\mathrm{SNR} = 4\eta N_p$. The BER is thus given by

$$\mathrm{BER} = \tfrac{1}{2}\,\mathrm{erfc}(\sqrt{\eta N_p/2}). \qquad \text{[ASK homodyne]} \tag{6.4.5}$$

Equations (6.4.4) and (6.4.5) can be used to calculate the receiver sensitivity at a specific BER. Similar to the direct-detection case discussed in Section 4.4, one can define the receiver sensitivity $\bar{P}_{\mathrm{rec}}$ as the average received power corresponding to a BER of 10^{-9}. From Eqs. (6.4.2) and (6.4.3), $\mathrm{BER} = 10^{-9}$ when $Q \approx 6$ or when $\mathrm{SNR} = 144$ (21.6 dB). For the ASK heterodyne case we can use Eq. (6.1.14) to relate SNR to $\bar{P}_{\mathrm{rec}}$ if we note that $\bar{P}_{\mathrm{rec}} = \bar{P}_s/2$ simply because signal power is zero during the 0 bits. The result is

$$\bar{P}_{\mathrm{rec}} = 2Q^2 h\nu\Delta f/\eta = 72 h\nu\Delta f/\eta. \tag{6.4.6}$$

For the ASK homodyne case, $\bar{P}_{\mathrm{rec}}$ is smaller by a factor of 2 because of the 3-dB homodyne-detection advantage discussed in Section 6.1.3. As an example, for a 1.55-μm ASK heterodyne receiver with $\eta = 0.8$ and $\Delta f = 1$ GHz, the receiver sensitivity is about 12 nW and reduces to 6 nW if homodyne detection is used.

The receiver sensitivity is often quoted in terms of the number of photons N_p by using Eqs. (6.4.4) and (6.4.5), as such a definition makes it independent of the receiver bandwidth and the operating wavelength. Furthermore, η is also set

to 1 so that the sensitivity corresponds to an ideal photodetector. It is easy to verify that for BER $= 10^{-9}$, $N_p = 72$ and 36 for the heterodyne and homodyne cases, respectively. It is important to remember that N_p corresponds to the number of photons within a single 1 bit. The average number of photons per bit, $\bar{N}_p$, is reduced by a factor of 2 if we assume that 0 and 1 bits are equally likely to occur in a long bit sequence, so that $\bar{N}_p = N_p/2$.

6.4.2 Synchronous PSK Receivers

Consider first the case of heterodyne detection. The signal at the decision circuit is given by Eq. (6.3.3) or by

$$I_d = \tfrac{1}{2}(I_p \cos\phi + i_c). \tag{6.4.7}$$

The main difference from the ASK case is that I_p is constant, but the phase ϕ takes values 0 or π depending on whether a 1 or 0 is transmitted. In both cases, I_d is a Gaussian random variable but its average value is either $I_p/2$ or $-I_p/2$, depending on the received bit. The situation is analogous to the ASK case with the difference that $I_0 = -I_1$ in place of being zero. In fact, one can use Eq. (6.4.2) for the BER, but Q is now given by

$$Q = \frac{I_1 - I_0}{\sigma_1 + \sigma_0} \approx \frac{2I_1}{2\sigma_1} = (\text{SNR})^{1/2}, \tag{6.4.8}$$

where $I_0 = -I_1$ and $\sigma_0 = \sigma_1$ was used. By using SNR $= 2\eta N_p$ from Eq. (6.1.15), the BER is given by

$$\text{BER} = \tfrac{1}{2}\,\text{erfc}(\sqrt{\eta N_p}). \qquad \text{[PSK heterodyne]} \tag{6.4.9}$$

As before, the SNR is improved by 3 dB, or by a factor of 2, in the case of PSK homodyne detection, so that

$$\text{BER} = \tfrac{1}{2}\,\text{erfc}(\sqrt{2\eta N_p}). \qquad \text{[PSK homodyne]} \tag{6.4.10}$$

The receiver sensitivity at a BER of 10^{-9} can be obtained by using $Q = 6$ and Eq. (6.1.14) for SNR. For the purpose of comparison, it is useful to express the receiver sensitivity in terms of the number of photons N_p. It is easy to verify that $N_p = 18$ and 9 for the cases of heterodyne and homodyne PSK detection, respectively. The average number of photons/bit, $\bar{N}_p$, equals N_p for the PSK format because the same power is transmitted during 1 and 0 bits. A PSK homodyne receiver is the most sensitive receiver, requiring only 9 photons/bit. It should be emphasized that this conclusion is based on the Gaussian approximation for the receiver noise [42].

It is interesting to compare the sensitivity of coherent receivers with that of a direct-detection receiver. Table 6.1 shows such a comparison. As discussed in Section 4.5.3, an ideal direct-detection receiver requires 10 photons/bit to operate at a BER of $\leq 10^{-9}$. This value is only slightly inferior to the best case of a PSK homodyne receiver and considerably superior to that of heterodyne

Table 6.1 Sensitivity of synchronous receivers

Modulation Format	Bit-Error Rate	N_p	$\bar{N}_p$
ASK heterodyne	$\frac{1}{2}\text{erfc}(\sqrt{\eta N_p/4})$	72	36
ASK homodyne	$\frac{1}{2}\text{erfc}(\sqrt{\eta N_p/2})$	36	18
PSK heterodyne	$\frac{1}{2}\text{erfc}(\sqrt{\eta N_p})$	18	18
PSK homodyne	$\frac{1}{2}\text{erfc}(\sqrt{2\eta N_p})$	9	9
FSK heterodyne	$\frac{1}{2}\text{erfc}(\sqrt{\eta N_p/2})$	36	36
Direct detection	$\frac{1}{2}\exp(-\eta N_p)$	20	10

schemes. However, it is never achieved in practice because of thermal noise, dark current, and many other factors, which degrade the sensitivity to the extent that $\bar{N}_p \approx 1000$ is usually required. In the case of coherent receivers, $\bar{N}_p$ below 100 can be realized simply because shot noise can be made dominant by increasing the local-oscillator power. The performance of coherent receivers is discussed in Section 6.6.

6.4.3 Synchronous FSK Receivers

Synchronous FSK receivers generally use a dual-filter scheme similar to that shown in Fig. 6.7(a) for the asynchronous case. Each filter passes only 1 or 0 bits. The scheme is equivalent to two complementary ASK heterodyne receivers operating in parallel. This feature can be used to calculate the BER of dual-filter synchronous FSK receivers. Indeed, one can use Eqs. (6.4.2) and (6.4.3) for the FSK case also. However, the SNR is improved by a factor of 2 compared with the ASK case. The improvement is due to the fact that whereas no power is received, on average, half the time for ASK receivers, the same amount of power is received all the time for FSK receivers. Hence the signal power is enhanced by a factor of 2, whereas the noise power remains the same if we assume the same receiver bandwidth in the two cases. By using SNR $= 4\eta N_p$ in Eq. (6.4.3), the BER is given by

$$\text{BER} = \tfrac{1}{2}\,\text{erfc}(\sqrt{\eta N_p/2}). \qquad \text{[FSK heterodyne]} \qquad (6.4.11)$$

The receiver sensitivity is obtained from Eq. (6.4.6) by replacing the factor of 72 by 36. In terms of the number of photons, the sensitivity is given by $N_p = 36$. The average number of photons/bit, $\bar{N}_p$, also equals 36, since each bit carries the same energy. A comparison of ASK and FSK heterodyne schemes in Table 6.1 shows that $\bar{N}_p = 36$ for both schemes. Therefore even though the ASK heterodyne receiver requires 72 photons within the 1 bit, the receiver sensitivity (average received power) is the same for both the ASK and FSK

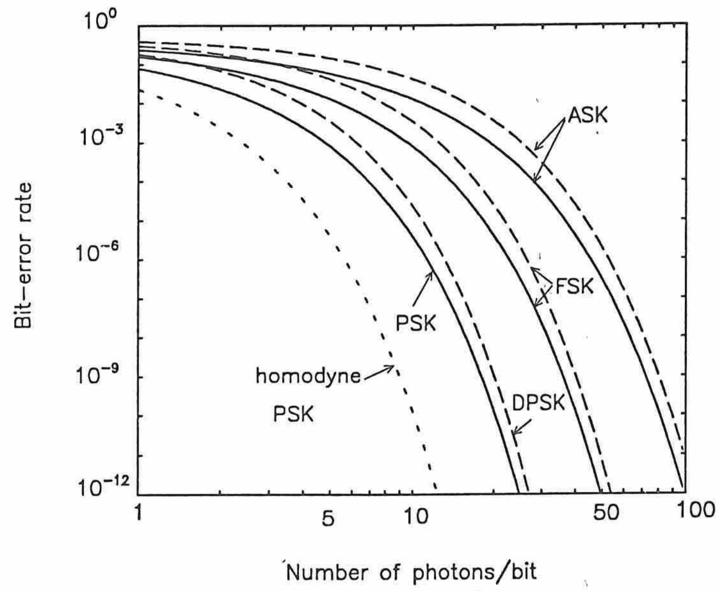

Figure 6.8 Bit-error-rate curves for various modulation formats. The solid and dashed lines correspond to the cases of synchronous and asynchronous demodulation, respectively.

schemes. Figure 6.8 plots the BER as a function of N_p for the ASK, PSK, and FSK formats by using Eqs. (6.4.4), (6.4.9), and (6.4.11). The dotted curve shows the BER for the case of synchronous PSK homodyne receiver discussed in Section 6.4.2. The dashed curves correspond to the case of asynchronous receivers discussed in the following subsections.

6.4.4 Asynchronous ASK Receivers

The BER calculation for asynchronous receivers is slightly more complicated than for synchronous receivers because the noise does not remain Gaussian when an envelope detector is used (see Fig. 6.6). The reason can be understood from Eq. (6.3.4), which shows the signal received by the decision circuit. For the case of an ideal ASK heterodyne receiver without phase fluctuations, ϕ can be set to zero so that (subscript d is dropped for simplicity of notation)

$$I = [(I_p + i_c)^2 + i_s^2]^{1/2}. \tag{6.4.12}$$

Even though both $I_p + i_c$ and i_s are Gaussian random variables, the probability density function (PDF) of I is not Gaussian. It can be calculated by using the standard techniques [43] and is found to be given by [44]

$$p(I, I_p) = \frac{I}{\sigma^2} \exp\left(-\frac{I^2 + I_p^2}{2\sigma^2}\right) I_0\left(\frac{I_p I}{\sigma^2}\right), \tag{6.4.13}$$

where I_0 represents the modified Bessel function of the first kind. Both i_c and i_s are assumed to have a Gaussian PDF with zero mean and the same standard deviation σ, where σ is the RMS noise current. The PDF given by Eq. (6.4.13) is known as the *Rice distribution* [44]. Note that I varies in the range 0 to ∞, since the output of an envelope detector can have only positive values. When $I_p = 0$, the Rice distribution reduces to the *Rayleigh distribution*, well known in statistical optics [43].

The BER calculation follows the analysis of Section 4.5.1 with the only difference that the Rice distribution be used in place of the Gaussian distribution. The BER is given by Eq. (4.5.2) or by

$$\text{BER} = \tfrac{1}{2}[P(0/1) + P(1/0)], \tag{6.4.14}$$

where

$$P(0/1) = \int_0^{I_D} p(I, I_1)\, dI, \qquad P(1/0) = \int_{I_D}^{\infty} P(I, I_0)\, dI. \tag{6.4.15}$$

The notation is the same as that of Section 4.5.1. In particular, I_D is the decision level and I_1 and I_0 are values of I_p for 1 and 0 bits. The noise is taken to be the same for all bits ($\sigma_0 = \sigma_1 = \sigma$) by assuming that it is dominated by the local-oscillator shot noise. The integrals in Eq. (6.4.15) can be expressed in terms of Marcum's Q function, defined as [45]

$$Q(\alpha, \beta) = \int_\beta^\infty x \exp\left(-\frac{x^2 + \alpha^2}{2}\right) I_0(\alpha x)\, dx. \tag{6.4.16}$$

The result for the BER is

$$\text{BER} = \frac{1}{2}\left[1 - Q\left(\frac{I_1}{\sigma}, \frac{I_D}{\sigma}\right) + Q\left(\frac{I_0}{\sigma}, \frac{I_D}{\sigma}\right)\right]. \tag{6.4.17}$$

The decision level I_D is chosen such that the BER is minimum for given values of I_1, I_0, and σ. It is difficult to obtain an analytic expression of I_D under general conditions. However, under typical operating conditions, $I_0 \approx 0$, $I_1/\sigma \gg 1$, and I_D is well approximated by $I_1/2$. The BER then becomes

$$\text{BER} \approx \tfrac{1}{2} \exp(-I_1^2/8\sigma^2) = \tfrac{1}{2} \exp(-\text{SNR}/8). \tag{6.4.18}$$

When the receiver noise σ is dominated by the shot noise, the SNR is given by Eq. (6.1.15). By using $\text{SNR} = 2\eta N_p$, we obtain the final result,

$$\text{BER} = \tfrac{1}{2} \exp(-\eta N_p/4), \tag{6.4.19}$$

which should be compared with Eq. (6.4.4) obtained for the case of synchronous ASK heterodyne receivers. Equation (6.4.19) is plotted in Fig. 6.8 with a dashed line. It shows that the BER is larger for the asynchronous case for the same value of ηN_p. However, the difference is so small that the receiver sensitivity at a BER of 10^{-9} is degraded by only about 0.5 dB. If we assume that $\eta = 1$, Eq. (6.4.19) shows that $\text{BER} = 10^{-9}$ for $N_p = 80$ ($N_p = 72$ for the synchronous case). Asynchronous receivers hence provide performance comparable to that of synchronous receivers and are often used in practice because of their simpler design.

6.4.5 Asynchronous FSK Receivers

Although a single-filter heterodyne receiver can be used for FSK, it has the disadvantage that one-half of the received power is rejected, resulting in an obvious 3-dB penalty. For this reason, a dual-filter FSK receiver [see Fig. 6.7(a)] is commonly employed in which 1 and 0 bits pass through separate filters. The output of two envelope detectors are subtracted, and the resulting signal is used by the decision circuit. Since the average current takes values I_p and $-I_p$ for 1 and 0 bits, the decision threshold is set in the middle ($I_D = 0$). Let I and I' be the currents generated in the upper and lower branches of the dual filter receiver, where both of them include noise currents through Eq. (6.4.12). Consider the case in which 1 bits are received in the upper branch. The current I is then given by Eq. (6.4.12) and follows a Rice distribution with $I_p = I_1$ in Eq. (6.4.13). On the other hand, I' consists only of noise and its distribution is obtained by setting $I_p = 0$ in Eq. (6.4.13). An error is made when $I' > I$, as the signal is then below the decision level, resulting in

$$P(0/1) = \int_0^\infty p(I, I_1) \left[\int_I^\infty p(I', 0)\, dI' \right] dI, \tag{6.4.20}$$

where the inner integral provides the error probability for a fixed value of I and the outer integral sums it over all possible values of I. The probability $P(1/0)$ can be obtained similarly. In fact, $P(1/0) = P(0/1)$ because of the symmetric nature of a dual-filter receiver.

The integral in Eq. (6.4.20) can be evaluated analytically. By using Eq. (6.4.13) in the inner integral with $I_p = 0$, it is easy to verify that

$$\int_I^\infty p(I', 0)\, dI' = \exp\left(-\frac{I^2}{2\sigma^2}\right). \tag{6.4.21}$$

By using Eqs. (6.4.14), (6.4.20), and (6.4.21) with $P(1/0) = P(0/1)$, the BER is given by

$$\mathrm{BER} = \int_0^\infty \frac{I}{\sigma^2} \exp\left(-\frac{I^2 + I_1^2}{2\sigma^2}\right) I_0\left(\frac{I_1 I}{\sigma^2}\right) \exp\left(-\frac{I^2}{2\sigma^2}\right) dI, \tag{6.4.22}$$

where $p(I, I_p)$ was substituted from Eq. (6.4.13). By introducing the variable $x = \sqrt{2}\, I$, Eq. (6.4.22) can be written as

$$\mathrm{BER} = \frac{1}{2} \exp\left(-\frac{I^2}{4\sigma^2}\right) \int_0^\infty \frac{x}{\sigma^2} \exp\left(-\frac{x^2 + I_1^2/2}{2\sigma^2}\right) I_0\left(\frac{I_1 x}{\sigma^2\sqrt{2}}\right) dx. \tag{6.4.23}$$

The integrand in Eq. (6.4.23) is just $p(x, I_1/\sqrt{2})$ and the integral must be 1. The BER is thus simply given by

$$\mathrm{BER} = \tfrac{1}{2}\exp(-I_1^2/4\sigma^2) = \tfrac{1}{2}\exp(-\mathrm{SNR}/4). \tag{6.4.24}$$

By using $\mathrm{SNR} = 2\eta N_p$ from Eq. (6.1.15), we obtain the final result

$$\mathrm{BER} = \tfrac{1}{2}\exp(-\eta N_p/2), \tag{6.4.25}$$

Table 6.2 Sensitivity of asynchronous receivers

Modulation Format	Bit-Error Rate	N_p	$\bar{N}_p$
ASK heterodyne	$\frac{1}{2}\exp(-\eta N_p/4)$	80	40
FSK heterodyne	$\frac{1}{2}\exp(-\eta N_p/2)$	40	40
DPSK heterodyne	$\frac{1}{2}\exp(-\eta N_p)$	20	20
Direct detection	$\frac{1}{2}\exp(-\eta N_p)$	20	10

which should be compared with Eq. (6.4.11) obtained for the case of synchronous FSK heterodyne receivers. Figure 6.8 compares the BER in the two cases. Just as in the ASK case, the BER is larger for asynchronous demodulation. However, the difference is small, and the receiver sensitivity is degraded by only about 0.5 dB compared with the synchronous case. If we assume that $\eta = 1$, $N_p = 40$ at a BER of 10^{-9} ($N_p = 36$ in the synchronous case). $\bar{N}_p$ also equals 40, since the same number of photons are received during 1 and 0 bits. Similar to the synchronous case, $\bar{N}_p$ is the same for both the ASK and FSK formats.

6.4.6 Asynchronous DPSK Receivers

As mentioned in Section 6.2.2, asynchronous demodulation cannot be used for PSK signals. A variant of PSK, known as DPSK, can be demodulated by using an asynchronous DPSK receiver [see Fig. 6.7(b)]. The filtered current is divided into two parts, and one part is delayed by exactly one bit period. The product of two currents contains information about the phase difference between the two neighboring bits and is used by the decision current to determine the bit pattern.

The BER calculation is slightly more complicated for the DPSK case, since the signal is formed by the product of two currents. The final result is, however, quite simple and is given by [11]

$$\text{BER} = \tfrac{1}{2}\exp(-\eta N_p). \tag{6.4.26}$$

It can be obtained from the FSK result, Eq. (6.4.24), by using a simple argument [13] which shows that the demodulated DPSK signal corresponds to the FSK case if we replace I_1 by $2I_1$ and σ^2 by $2\sigma^2$. Figure 6.8 shows the BER by a dashed line (the curve marked DPSK). For $\eta = 1$, a BER of 10^{-9} is obtained for $N_p = 20$. Thus DPSK is more sensitive by 3 dB compared with both ASK and FSK. Table 6.2 lists the BER and the receiver sensitivity for the three modulation schemes used with asynchronous demodulation. The direct-detection case is also listed for comparison.

Synchronous Modulators

Solitons can also be controlled in the time domain. A 1991 experiment demonstrated [93] transmission of a soliton train over 1 million kilometers using the technique of *synchronous intensity modulation*, implemented using a $LiNbO_3$ modulator. The technique works by introducing additional losses for solitons that have shifted from their original position (center of the bit slot). The modulator forces solitons to move toward its transmission peak where the loss is minimum. Mathematically, the action of modulator is to change the soliton amplitude as

$$u_s(\xi_m, \tau) \to T_m(\tau - \tau_m) u_s(\xi_m, \tau), \qquad (10.3.27)$$

where $T_m(\tau)$ is the transmission coefficient of the modulator located at $\xi = \xi_m$. For the case of sinusoidal modulation, commonly used in practice, $T_m(\tau)$ can be approximated by a parabola in the vicinity of $\tau = \tau_m$. The soliton perturbation theory can again be used in a manner similar to the case of optical filters. The results show that synchronous modulators force the soliton to move toward their transmission peak, and such forcing reduces the timing jitter considerably.

Synchronous modulation can be combined with optical filters to control solitons simultaneously in both the time and frequency domains. A numerical study [94], followed by an experimental realization [95], indicated the possibility of achieving arbitrarily long transmission distances by this combination. Synchronous intensity modulation also permits a relatively large amplifier spacing [96] because it can reduce the impact of dispersive waves. This property of modulators has been exploited to transmit a 20-Gb/s soliton train over 150,000 km with an amplifier spacing of 105 km [97]. In another experiment, a single synchronous modulator, inserted just after the transmitter, allowed transmission of a 20-Gb/s signal over 3000 km [98], well beyond the 2300-km Gordon–Haus limit of the system without modulation. In this experiment, the clock signal used to generate the soliton train was also used to drive the modulator. In contrast, when in-line synchronous modulators are used, the signal clock must be regenerated electronically.

Synchronous modulation can also be implemented by using a phase modulator [99], [100]. One can understand the effect of periodic phase modulation by recalling that a frequency shift $\delta\omega = -\partial\phi(t)/\partial t$ is associated with any phase variation $\phi(t)$. Since a change in soliton frequency is equivalent to a change in the group velocity, phase modulation induces a temporal displacement. Synchronous phase modulation is implemented in such a way that the soliton experiences a frequency shift only if it moves away from the center of the bit slot, thereby confining it to its original position despite the timing jitter induced by ASE and other sources. Intensity and phase modulations can be combined together [101] to further improve the system performance. Similar to the case of sliding-frequency optical filters, synchronous modulators help a soliton communication system in several other ways. Among other things, they reduce soliton interaction, clamp the level of amplifier noise, and inhibit the growth of dispersive waves [102].